

Comments on “Reanalyses and Observations: What’s the Difference?”

—ROSS N. HOFFMAN

NOAA/Atlantic Oceanographic and Meteorological Laboratory,
and Cooperative Institute for Marine and Atmospheric Studies,
University of Miami, Miami, Florida

—NIKKI PRIVÉ

Morgan State University, Baltimore, Maryland

—MARK BOURASSA

The Florida State University, Tallahassee, Florida

In a very readable essay, Parker (2016) highlights the similarities between reanalyses and observations. Parker’s essay is thought provoking, but there are important omissions. Parker makes the point that all geophysical data can be thought of as “measurement outcomes.” However, in practice, it is not useful to consider all geophysical data to be of the same nature, with differences being a matter of degree only. Indeed, Parker differentiates several different types of geophysical data, which will be reprised below. Parker concludes that the uncertainty of reanalyses is more complex and less understood and less quantified than the uncertainty of observations. The corollary to this conclusion, and the theme of this essay, cannot be stressed enough: *it is imperative that researchers understand the sources, uncertainty, biases, and other limitations of any data that they use.*

In a concise essay, all aspects of the question “What’s the difference?” could not be covered. Here, Parker’s discussion is amended and expanded. First, readers of Parker’s essay should be aware that for many purposes, measurements, retrievals, and

analyses are not interchangeable and should be treated differently.¹ In this essay, a “measurement” is a direct traceable observation of some geophysical quantity; a “retrieval” is a combination of measurements (e.g., radiances from a satellite) and prior information and is an indirect observation or derived measurement in the terminology of Parker; and an “analysis” is the result of a data assimilation (DA) or other interpolative process that combines diverse observations and a background or prior, normally a short-range forecast.² As used here, measurements and retrievals are “observations,” and observations and analyses are “data.” Second, different scales and different quantities are observed or represented by an in situ sensor (e.g., a temperature measured by a sensor on a radiosonde), a satellite sensor [e.g., a retrieved Atmospheric Infrared Sounder (AIRS) temperature], and an analysis [e.g., the temperature at a grid location in a European Centre for Medium-Range Weather Forecasts (ECMWF) operational analysis]. Third, depending on one’s purpose, the scales for validating geophysical data may be different, and hence error characterization could depend on the user’s goals. Fourth, there are important limitations of analyses. Observations are irregular in space and time, analyses are not, but at a cost: in situations where the observations are lacking, the analysis procedure relies on imperfect statistical and forecast model information. These limitations are accounted for in well-conducted research. This essay will expand these points, emphasizing key aspects of data that are often overlooked and can impact the suitability of data for a specific application.

THE RELATIONSHIP OF GEOPHYSICAL DATA TYPES TO REALITY. A useful analogy is to think of geophysical measurements as fossils—the imperfect imprints of reality preserved by a variety of more or less reliable mechanisms. In this analogy,

¹ A reanalysis is a special type of analysis, and the “re” will be dropped when the discussion applies to both analyses and reanalyses.

² Data assimilation differs from other interpolative processes in that the prior is the forecast from the previous analysis.

a retrieval is a skeleton in a museum with some parts reconstructed based on principles of general anatomy, and an analysis is a computer-graphics-generated animation—the depiction of reality based in part on fossil evidence and in part by physics simulation. A fossil is several steps removed from a dinosaur, and an animation, no matter how “realistic,” is even further removed. This analogy only goes so far, but it sets the stage for the following discussion of the ways in which geophysical data of various types are abstractions of reality—something all users should keep in mind.

First, and foremost, an analysis is fundamentally a “model” of the atmosphere, that is, a quantitative yet simplified representation of the atmosphere in reality [see Rosenblueth and Wiener (1945) for an in-depth discussion of the concept of models].³ There are many possible objectives for a reanalysis, including the understanding of atmospheric processes, the estimation of various statistics of the atmosphere, and so forth.⁴

In practice, in all models some elements of the actual “thing” are abstracted or mapped into the model. For an analysis, a principal abstraction is discretization, which results in reducing (eliminating) the information about the smaller (smallest) scales in reality.

Parker notes that observations may involve some modeling in the process of converting the raw measurements into the final observations, or may be used to develop a model. One could go further to say that as soon as an observation is put to any use in representing reality, that observation itself becomes a model.

For the important example of satellite infrared and microwave sensors, the instrument is engineered to measure radiance; however, the actual measurement might be photon counts, which must be converted to radiances. This involves calibration, but according to Wielicki et al. (2013) the conversion is in principle traceable to the International System of Units (Système International d’Unités or SI). On this basis, radiances are considered here to be measurements. Note that most modern DA systems assimilate radiances, not retrievals. Radiances are often referred to as a level 1 or sensor data record (SDR) product, while retrievals are often referred to as a level 2 or an environmental data record (EDR) product. When EDRs are binned or analyzed on a horizontal grid, the result may be termed a level 3

product. Level 3 products include varying degrees of prior information and may be considered analyses.

REPRESENTATIVENESS. Geophysical data differ in what processes and what scales are represented. This is a critical consideration for users of the data. Parker discusses some of these differences, but not the basic and critical differences between the spatial and temporal scales of analyses and observations. There is only one real atmosphere, but each analysis or observation inevitably filters reality to match the scales representable by the analysis or observation.

In general, given the difference between two different data types, representativeness error is the component of that difference that arises from spatial and temporal scales represented by one, but not the other, type of data. Validation studies of satellite sensors provide valuable insights about representativeness issues that may arise in using geophysical data. For example, when using ship observations to validate satellite winds, which have a sampling footprint of approximately 25 km, it is important to average the ship observations in time to filter the small scales, with the averaging interval increasing with decreasing wind speed in accord with Taylor’s frozen turbulence hypothesis that equates temporal and spatial variability (Bourassa et al. 2003). Such a trade-off (of temporal for spatial variability) may not be sufficient when the sources of variability are inhomogeneous. For example, in their discussion of sea surface salinity (SSS) validation, Boutin et al. (2016) note that the “spatiotemporal variability of SSS within a satellite footprint (50–150 km) is a major issue for satellite SSS validation in the vicinity of river plumes, frontal zones, and significant precipitation.”

Representativeness is related to the specification of uncertainty in DA systems: in the DA context, representativeness error is the variability present in observations, but not represented by the DA system, and is considered a component of observation error.⁵ This, of course, is a DA-centric viewpoint, but is consistent with the DA process, which seeks the optimal fit to observations and prior information within the space of feasible solutions, that is, representable by the forecast model that is used. In many practical cases, representativeness errors dominate all other error sources combined. Note that from this DA-centric viewpoint, an analysis is expected to have smaller errors than observations on the scales represented by the analysis. Artifacts in

³ In this discussion the word “model” is reserved for this generic concept, which should not be confused with the term “forecast model.”

⁴ For clarity, this essay focuses on the atmosphere, but many of the general statements are applicable to other geophysical systems.

⁵ Sometimes, but not in this discussion, representativeness error is defined to include both forward model (i.e., simulation) error and differences in scales represented.

analyses can occur when the representativeness error is inaccurately estimated (e.g., Smith et al. 2011).

Another result of the DA-centric viewpoint is that not all the information present in observations is assimilated into the analysis. In DA systems, dense satellite observations are often averaged into so-called “superobservations” that are more consistent with the scales of the DA system. For example, Lin et al. (2016) report that, in the ECMWF DA system, assimilating superobservations of satellite wind data at a resolution of 50–100 km is more effective than assimilating the original 25-km product. But such decisions trading off resolution and representativeness error for DA purposes will impact both the noise and the representation of small scales in estimates of the curl of wind stress, critical for oceanographic applications (Collins et al. 2012; Holbach and Bourassa 2017).

When using observations, especially remotely sensed observations, it is not just horizontal resolution that is important, but also, as Parker noted, a precise specification of just what is being measured. This is especially so at ocean and land surfaces, as the following examples show. Scatterometers do not actually measure the 10-m wind directly, but rather the reflectivity of the surface to the transmitted radar signal, which is empirically related to 10-m neutral stability wind (Kara et al. 2008; Wentz et al. 2017). Passive microwave radiometers do not actually measure quantities like surface temperature, but rather the apparent brightness temperature of the surface as seen through the atmosphere. For example, for microwave sensing of soil temperature and moisture, both of which have diurnally varying boundary conditions, longer wavelength channels respond to deeper layers (Entekhabi et al. 1994; Moncet et al. 2011; Galantowicz et al. 2011). As another example, in the ocean there can be great variations in temperature and salinity just below the surface, and different observing methodologies effectively sample different depths (Donlon et al. 2007; Boutin et al. 2016). These details are critical when such data are assimilated into coupled DA systems or used to characterize fluxes between land and atmosphere and ocean and atmosphere. However, such processes are often highly parameterized (i.e., not actually resolved) by land or ocean forecast models, in part because the vertical scale of the process in reality is so much smaller than the vertical discretization of the forecast model.

THE UNCERTAINTY OF UNCERTAINTY.

Geophysical data should only be used in ways consistent with the data uncertainty. The addition of prior information in an ill-posed retrieval or analysis

problem renders the problem well posed. The quality of the prior information has a direct impact on the quality of the retrieval or analysis. Because of the use of a forecast in the analysis, the characterization of analysis uncertainty is complex. In contrast, for retrievals, the estimated errors are usually well defined and, for well-posed retrievals, are quite small for the spatial/temporal scale represented by the observations (e.g., Wentz et al. 2017).

Parker advocates the inclusion of uncertainty estimates along with reanalysis datasets. While this would appear to be a good suggestion on the face of it, there are some complications to quantifying analysis uncertainties. As a result, providing “one size fits all” uncertainty metrics might mislead users into assuming that the published uncertainties are valid for all applications. There are two types of meta-uncertainty that interact. First, there is uncertainty in mapping the analysis to the phenomena of interest in reality. As a model, the analysis fields may (or may not) have a precise definition in relation to the state of the real world. For example, the temperature field in an analysis may be explicitly defined as some weighted spatiotemporal average of temperatures over a volume, or such an explicit definition may be omitted. The uncertainty of the analysis in comparison to reality is a function of the definition of each analysis field. Regardless of how the analysis fields may be defined, each user may have a different application of the fields with a correspondingly different measure of uncertainty. For example, a user trying to validate a climate model wind field and a user interested in evaluating a location for wind energy may be interested in very different statistical aspects of the same wind field, with correspondingly different measures of uncertainty. As a result, users should consider, in the context of their goals, the way in which they interpret the analysis, and how that interpretation relates to reality.

Second, there is uncertainty in specifying the uncertainty of the analysis. Actual uncertainty (or accuracy) of analyses varies among DA systems (e.g., Peña and Toth 2014). Further, observing networks evolve over time and forecast model error varies with season and location. Therefore, uncertainty for a given analysis varies in time and space (e.g., Feng et al. 2017). To properly report uncertainties, four-dimensional fields should be developed for each variable.

In fact, a proper characterization of analysis uncertainty should go beyond standard deviations in four dimensions for each variable. Modern ensemble DA systems produce ensemble representations of the uncertainty that are not constrained, except by ensemble size. On the other hand, providing instead a reduced dataset

of uncertainties might lead to uncertainty ranges that are unhelpful or misleading. For example, the wind energy user might be interested in the uncertainty of kinetic energy integrated over a specific volume and the correlation of this quantity from location to location. This is a straightforward calculation for an ensemble of analysis. However, if standard deviations are the only available measure of uncertainty, then this calculation requires difficult-to-justify assumptions about the structure of the wind field. Data access tools should be extended to help users in mapping analysis ensemble uncertainty to user-defined uncertainty metrics.

LIMITATIONS OF ANALYSES. The large number of studies, which call into question the ability of different analyses to represent particular phenomena, should be a warning signal to all users of analyses. (Of course, observations also misrepresent geophysical phenomena, due to accuracy, representativeness, and coverage limitations.) In the cases listed below, the investigators attempted to validate the use of analyses for particular phenomena by comparison to observation datasets that properly represented the phenomena of interest, but with limited coverage. If the analyses could be validated, they would provide a much more comprehensive dataset for the study of the phenomena. However, in these cases the phenomena of interest are not properly represented by the analysis. Consequently, the analyses have spatially coherent and correlated errors, which may not be properly captured by estimates of analysis uncertainty. While observations may have correlated errors, the structure of analysis errors in cases such as these can be complex.

The following list, chosen to show a diversity of phenomena, is just a sample:

- For the equatorial lower stratosphere, Podglajen et al. (2014) find reanalyses misrepresent certain types of large-scale motions (specifically, equatorial Kelvin and Yanai wave packets).
- For polar lows, Laffineur et al. (2014) and Zappa et al. (2014) find that reanalyses detected only about half of the observed polar lows—small-scale hurricane-like storms found at high northern latitudes. In particular, the smaller-scale storms are missing (Condron et al. 2008).
- For the vertical structure of the lower atmosphere, Serreze et al. (2012) find that low-level temperature inversions are not well captured by reanalyses.
- For the energy and water cycles, Rienecker et al. (2011) find that precipitation and fluxes are not well constrained in reanalyses.

- For marine surface winds, Li et al. (2013) find that all reanalyses are too conservative, with large positive speed biases for weak winds and large negative speed biases for strong winds.
- For dust lifting, Largeron et al. (2015) find that reanalyses do not properly represent the Sahelian surface winds.

These and other studies of the same type reinforce the concern that generic issues with DA systems—lack of sufficient observations, limited resolution, incorrect specification of error statistics, and imperfect forecast models—are cause for the user to be wary of equating analyses with observations.

CONCLUDING REMARKS. Analyses and reanalyses are not universally applicable. Analyses are particularly useful when validating forecast or climate model products, as there are many more commonalities than with observations—time and space are discrete, the effects of physical processes on the tendencies of variables are parameterized, etc. But even in this case, the utility of the analyses is constrained by limited knowledge of the associated uncertainty. In general, a researcher relying on an analysis as a model of reality should be cautious. Similarly, observations can easily be misused if the user is unfamiliar with their characteristics as well as their strengths and weaknesses. Users of analyses (and observations) should familiarize themselves with technical documents and publications that describe and evaluate the analysis quality, or undertake validation themselves, and make an effort to understand the trustworthiness of the analysis for their specific purpose. In conclusion, geophysical data are diverse. Know your data! Do not use data beyond their limitations.

ACKNOWLEDGMENTS. The authors thank Parker for her original essay and her reply to this comment. In spite of the length of this comment, the authors agree with Parker on almost all issues discussed in this exchange. The authors also thank several colleagues and the reviewers for their comments, suggestions, and encouragement.

REFERENCES

- Bourassa, M. A., D. M. Legler, J. J. O'Brien, and S. R. Smith, 2003: SeaWinds validation with research vessels. *J. Geophys. Res.*, **108**, 3019, doi:10.1029/2001JC001028.
- Boutin, J., and Coauthors, 2016: Satellite and in situ salinity: Understanding near-surface stratification and subfootprint variability. *Bull. Amer. Meteor. Soc.*, **97**, 1391–1407, doi:10.1175/BAMS-D-15-00032.1.

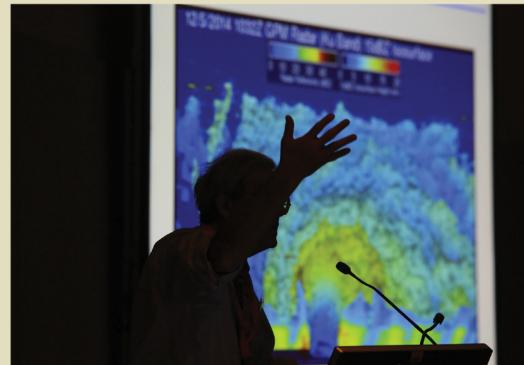
- Collins, C., C. J. C. Reason, and J. C. Hermes, 2012: Scatterometer and reanalysis wind products over the western tropical Indian Ocean. *J. Geophys. Res.*, **117**, C03045, doi:10.1029/2011JC007531.
- Condron, A., G. R. Bigg, and I. A. Renfrew, 2008: Modeling the impact of polar mesocyclones on ocean circulation. *J. Geophys. Res.*, **113**, C10005, doi:10.1029/2007JC004599.
- Donlon, C., and Coauthors, 2007: The Global Ocean Data Assimilation Experiment High-Resolution Sea Surface Temperature Pilot Project. *Bull. Amer. Meteor. Soc.*, **88**, 1197–1213, doi:10.1175/BAMS-88-8-1197.
- Entekhabi, D., H. Nakamura, and E. G. Njoku, 1994: Solving the inverse-problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. *IEEE Trans. Geosci. Remote Sens.*, **32**, 438–448, doi:10.1109/36.295058.
- Feng, J., Z. Toth, and M. Peña, 2017: Spatially extended estimates of analysis and short-range forecast error variances. *Tellus*, **69A**, 1325301, doi:10.1080/16000870.2017.1325301.
- Galantowicz, J. F., J.-L. Moncet, P. Liang, A. E. Lipton, G. Uymin, C. Prigent, and C. Grassotti, 2011: Subsurface emission effects in AMSR-E measurements: Implications for land surface microwave emissivity retrieval. *J. Geophys. Res.*, **116**, D17105, doi:10.1029/2010JD015431.
- Holbach, H. M., and M. A. Bourassa, 2017: Platform and across-swath comparison of vorticity spectra from QuikSCAT, ASCAT-A, OSCAT2, and ASCAT-B scatterometers. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **10**, 2205–2213, doi:10.1109/JSTARS.2016.2642583.
- Kara, A. B., A. J. Wallcraft, and M. A. Bourassa, 2008: Air-sea stability effects on the 10 m winds over the global ocean: Evaluations of air-sea flux algorithms. *J. Geophys. Res.*, **113**, C04009, doi:10.1029/2007JC004324.
- Laffineur, T., C. Claud, J.-P. Chaboureau, and G. Noer, 2014: Polar lows over the Nordic Seas: Improved representation in ERA-Interim compared to ERA-40 and the impact on downscaled simulations. *Mon. Wea. Rev.*, **142**, 2271–2289, doi:10.1175/MWR-D-13-00171.1.
- Largergeron, Y., F. Guichard, D. Bouniol, F. Couvreur, L. Kergoat, and B. Marticorena, 2015: Can we use surface wind fields from meteorological reanalyses for Sahelian dust emission simulations? *Geophys. Res. Lett.*, **42**, 2490–2499, doi:10.1002/2014GL062938.
- Li, M., J. Liu, Z. Wang, H. Wang, Z. Zhang, L. Zhang, and Q. Yang, 2013: Assessment of sea surface wind from NWP reanalyses and satellites in the Southern Ocean. *J. Atmos. Oceanic Technol.*, **30**, 1842–1853, doi:10.1175/JTECH-D-12-00240.1.
- Lin, W., G. de Chiara, M. Portabella, A. Stoffelen, J. Vogelzang, and A. Verhoef, 2016: On the assimilation of ASCAT winds. *Proc. 2016 IEEE Int. Geoscience and Remote Sensing Symp.*, Beijing, China, IEEE, 2269–2271, doi:10.1109/IGARSS.2016.7729586.
- Moncet, J.-L., P. Liang, J. F. Galantowicz, A. E. Lipton, G. Uymin, C. Prigent, and C. Grassotti, 2011: Land surface microwave emissivities derived from AMSR-E and MODIS measurements with advanced quality control. *J. Geophys. Res.*, **116**, D16104, doi:10.1029/2010JD015429.
- Parker, W. S., 2016: Reanalyses and observations: What's the difference? *Bull. Amer. Meteor. Soc.*, **97**, 1565–1572, doi:10.1175/BAMS-D-14-00226.1.
- Peña, M., and Z. Toth, 2014: Estimation of analysis and forecast error variances. *Tellus*, **66A**, 21767, doi:10.3402/tellusa.v66.21767.
- Podglajen, A., A. Hertzog, R. Plougonven, and N. Žagar, 2014: Assessment of the accuracy of (re)analyses in the equatorial lower stratosphere. *J. Geophys. Res. Atmos.*, **119**, 11 166–11 188, doi:10.1002/2014JD021849.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.
- Rosenblueth, A., and N. Wiener, 1945: The role of models in science. *Philos. Sci.*, **12**, 316–321, doi:10.1086/286874.
- Serreze, M. C., A. P. Barrett, and J. Stroeve, 2012: Recent changes in tropospheric water vapor over the Arctic as assessed from radiosondes and atmospheric reanalyses. *J. Geophys. Res.*, **117**, D10104, doi:10.1029/2011JD017421.
- Smith, S. R., P. J. Hughes, and M. A. Bourassa, 2011: A comparison of nine monthly air-sea flux products. *Int. J. Climatol.*, **31**, 1002–1027, doi:10.1002/joc.2225.
- Wentz, F. J., and Coauthors, 2017: Evaluating and extending the ocean wind climate data record. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **10**, 2165–2185, doi:10.1109/JSTARS.2016.2643641.
- Wielicki, B. A., and Coauthors, 2013: Achieving climate change absolute accuracy in orbit. *Bull. Amer. Meteor. Soc.*, **94**, 1519–1539, doi:10.1175/BAMS-D-12-00149.1.
- Zappa, G., L. Shaffrey, and K. Hodges, 2014: Can polar lows be objectively identified and tracked in the ECMWF operational analysis and the ERA-Interim reanalysis? *Mon. Wea. Rev.*, **142**, 2596–2608, doi:10.1175/MWR-D-14-00064.1.

AMS membership

RENEWALS ARE OPEN!

DON'T MISS OUT!

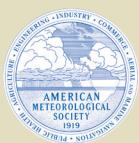
AMS membership provides value throughout the year and keeps you connected to the weather, water, and climate community. Your work is vital to furthering scientific understanding and AMS helps ensure it makes a difference.



Your AMS Membership provides:

- Access to a vibrant community of peers and colleagues
- Discounts on meetings registrations
- Subscriptions to *BAMS* and *Physics Today*
- Access to top scientific journals
- Special privileges with the AMS online Career Center

Renew your membership today at
ametsoc.org/renew



AMS
American Meteorological Society